DOCUMENT RESUME

ED 111 024 CS 202 243

AUTHOR Follettie, Joseph F.
TITLE On Size of the Lexical Domain for Primary Education.
INSTITUTION Southwest Regional Laboratory for Educational Research and Development, Los Alamitos, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
REPORT NO SWRL-TM-1-71-7
PUB DATE 71
NOTE 19p.

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS Concept Formation; *Dictionaries; *Fundamental Concepts; *Lexicology; Primary Education; Reading Research

ABSTRACT

The specification of a lexicon, centering on the referential base concept and compiled for primary education, is explored in this paper. Classification of a sample of 518 entries from a desk dictionary into one of five categories is discussed, and analysis of the 266 entries in the fourth category (base concept-naming entries) according to the base word and its associated entries is illustrated in a table. (JM)

SOUTHWEST REGIONAL LABORATORY
TECHNICAL MEMORANDUM

DATE: June 7, 1971

NO: TM-1-71-7

ON SIZE OF THE LEXICAL DOMAIN FOR PRIMARY EDUCATION

Joseph F. Follettie

ABSTRACT

How specification of a lexicon for primary education might be related to the notion of the referential base concept is illustrated using a sample of 518 entries from a desk dictionary.

ON SIZE OF THE LEXICAL DOMAIN FOR PRIMARY EDUCATION

The size of primary education's lexicon will depend on two factors: a) how many concepts and special phenomena are to be taught and b) how the lexicon relates to such referents. The first factor--extent of the content domain for primary education--presently can be appreciated only intuitively in terms of a hypothetical mean acquisition rate. The second--how the collection of names of phenomena of interest relate to what these names signify--cannot even presently be appreciated at a useful intuitive level. Hence, it would be premature to try to deal definitively with size of primary education's lexicon at this time. This paper will explore how an objective of lexical specification might be pursued where interest centers on concept learning.

We assume a lexical universe, which provides the names of all concepts and special phenomena that are of interest to mankind at any given time, and a variety of referential universes into which classify the named concepts and special phenomena. Were the concepts and special phenomena of a given referential universe arbitrary collections, then the naming of these phenomena could be straightforwardly simple. However, most referential universes have a complex multidimensional relational structure; the lexicon, as constructed and formally interpreted using the conventions of linguistics, attempts to mirror the relational complexity of referential universes. In consequence, an array of lexical items, taken together with their definitions and listings of synonyms, constitutes a calculus of signification, however crude.

Two general approaches to relating the lexical universe to referential universes by linguistic means have evolved. These are the dictionary and thesaurus approaches.

1. The dictionary approach takes the lexical items as its point of departure. Given a specified lexical item, one uses an alphabetically-organized dictionary to locate the item and to obtain a linguistic characterization or set of characterizations of the item's referential signification (or meaning).

2. The thesaurus approach takes a referential signification as its point of departure. A general thesaurus organizes entries on a subject-matter basis, rather than alphabetically. Like a dictionary, a thesaurus deals with a large number of definienda. Unlike a dictionary, these definienda are organized on a subject-matter basis and are characterized by a field of lexical entries that are synonyms or approach being synonym's in the definiendum's universe of discourse.

Over a decade ago the Cambridge Group found a thesaurus approach useful to the formulation of computer-based information retrieval system designs. Later papers dealing with the question of multiple meanings of

lexical entries might find it useful to reference the analysis of lexicon more to a thesaurus than to a dictionary. However, the multiple meaning question will not be systemically addressed here.

A good desk dictionary offers many clues to the identity of concepts that lexicographers believe some members of the general population will find useful. A user equipped with a not-inordinate number of interpreted base concept names and a reasonable understanding of the linguistic system can obtain a liberal education just by studying a desk dictionary. The assumption will be entertained that a desk dictionary can be used to identify all of the general and much of the technical lexicon for English and to differentiate this lexicon into primary and postprimary educational components. The desk dictionary that will be used for this purpose is Webster's Seventh New Collegiate Dictionary. (Only the main body of Webster's will be used; the appended sections--e.g., Bibliographic Names, Pronouncing Gazetteer--will not be considered.) A sample of the lexical entries of Webster's will be used to estimate the number of concepts that Webster's names and the number of concepts in each of certain subclassifications of interest.

A count of the entries on every 10th page of Webster's--52 of 1041 pages--suggests that the dictionary contains approximately 70,000 (70K) entries. Some of these entries--e.g., CYBELE: a nature goddess of the ancient peoples of Asia Minor--refer to a specific thing, place, state, or process, rather than to a concept. Some--e.g., ^{5}LONG: to be suitable or fitting--are archaic. Some--e.g., HIRPLE: limp, hobble--use a concept name that occurs only in a non-American dialect of English. Some--e.g., PSEUDEPIGRAPHY: the ascription of false names of authors to works--do not seem compelling candidates for interpretation in a primary educational setting. A review of the range of entries contained in Webster's suggests that purposes of the paper will be served by classifying entries into one of five mutually-exclusive categories. These categories are:

Category 1. This category embraces non-concept entries, archaic entries, entries inappropriate to those who speak an American dialect of English, and entries referring to concepts that are sufficiently specialized to warrant an assumption that they probably will not be treated during primary education.

Category 2. Here are included entries that are multiple-root or multiple-word and appear to warrant the assumption that they probably will be treated during primary education. Subcategories are a) compound words, b) hyphenated words, c) phrases, and d) multiple-root words at least one root of which is a combining form. While Category 2 entries are viewed as naming "higher level" concepts on the basis of having more than one root, some--particularly certain of the compound words--could turn out to be "lower level" when categorized on a meaning basis--e.g., BLUEBIRD, CATFISH.

Category 3. This category embraces combining forms and affixes. A combining form is considered to reflect a root element and an affix not to. The thinness of this distinction, when viewed from a meaning standpoint, is pointed up by the prefix entries UNI- and BI-, signifying one and two, which contrast with the combining form entries MONO- and DI-, signifying single and double. Entries classified as affixes for the most part name concepts to be taught during primary education. Those classified as combining forms may or may not be taught during primary education--some will and some will not.

Category 4. It is assumed that Category 4 entries surely will be treated during primary education. Here are included the dictionary's "base concept-naming entries" and those extensions, through affixation, of base entry forms that will be understood in consequence of word formation instruction featuring linguistic rules, to be administered during primary education.

Category 5. This category embraces entries that are repetitive of Category 4 entries or entries that are extensions of a base form of Category 4. The category exists only to serve purposes of sample analysis.

A Category 4 entry E_i has a set of associated entries A_i, where the set $(E_i + A_i)$ represents purely-linguistic extensions of a base form B_i. By "purely-linguistic" is only meant that if a base form--e.g., compose--and the significance of certain linguistic rules--e.g., $N_i = V_i$ + ition--are understood, then an extension of a base form--e.g., COMPOSITION--will be understood. That is, no extension of a base form will be allowed that involves a denotative shift from the base form that cannot be accounted for by an appropriate linguistic rule.

Even where preservation of meaning is required, state-of-the-art probably is consonant with a more-satisfactory base form specification than has been attempted below. The intuitive approach to base form specification adopted here probably will suffice for purposes of preliminary analysis. (Table 2, to be discussed later, illustrates the nuances of base form specification as practiced in the present analysis.)

The first step of the analysis to be reported addressed the distribution of entries across the five categories described above. An estimate of the proportion of entries falling into each category was based on a sample of 518 entries drawn from Webster's. For all even pages of Webster's--there are 520--the 4th entry down of the 2nd column was drawn. Two pages contained no such entry. Hence, the sample contains 518 entries. Based on the defining information accompanying the entry, it was classified into one of the five categories. Whether the entry should be considered to belong to a postprimary lexicon (Category 1) or to a primary lexicon (Categories 2, 4) was decided on a purely intuitive basis.

Table 1 summarizes categorization findings. The number of entries estimated to fall in each category was obtained by multiplying the category's proportion of sample entries by 70K, the estimate of number of entries in Webster's.

Table 1.

Distribution of the Sample of 518 Webster's Entries across Categories

Category	No. of Entries	Percent Of Sample	Estimated Size of Category Universe
1	131	25.3	18K
2	107	20.7	14K
3	13	2.5	2K
4	266	51.4	36K
5	1	0.2	---

Classification into Categories 3 and 5 is straightforward. A degree of reclassification involving entries of Categories 1, 2, and 4 is to be expected. The data would not prove useful to devising a taxonomy and evaluating some of its implications only if there were a bias toward misclassification into one of the categories more than toward misclassification into the others. Category 4 entries (together with apparent base forms and associated entries that are extensions of the base form) are presented in Table 2. Entries for the other categories are presented, by category, in Table 3. Intuition suggests that most adults will not have most of the concepts that are named by Category 1 entries and that many children in junior high school will have many of the concepts that are named by Category 2 and Category 4 entries. While not all of the entries of the sample probably are definitively classified, there is little reason to believe that categorical proportionality, as reflected in Table 1, will not prove useful for purposes of making preliminary projections referenced to Webster's.

Table 2.

Category 4 Entries in the Webster's Sample of 518 Entries[a]

Entry(E)	Base(B)	Associated Entries(A)	Sum A
abjuration	abjure	B	1
abundance	abound	B, abundant	2
accordingly	accord	^{1}B, ^{2}B, ^{1}disB, Bance, Bant ^{2}disB	6
acorn			
^{2}adagio			
admiralty			
afghan			
align	align	Bment	1
amour	amor-	Bous, Bist	2
angel			
annotation	annotate	B	1
anthropic	anthrope	misB, misBic, misBy	3
^{3}any	any	^{1}B, ^{2}B	2
artillery	artiller-	Bist	1
aspheric	sphere	^{1}B, ^{2}B, Bal, Bical, Boid, Boidal, Bule, By	8
astigmatism	astigmat-	Bic	1
atomization	atom	B, Bic, Bicity, Bism, Bistic, Bize, Bizer	7
axilla	axill-	^{1}Bary, ^{2}Bary, Bar	3
balk			
baronage	baron	B, Bess, Bial, By	4
bassoon			
belittle	little	^{1}B, ^{2}B	2
^{5}blow			
^{3}bossy	boss	^{5}B, Biness	2
^{1}box			
cadet			
caliber	calib(e)r-	Bate	1
camise			
capricious	caprice	B	1
charisma			

[a]Superscripts are those used by Webster's to distinguish different entries spelled alike. Conjugational and other forms appearing in bold face in an entry's defining information are excluded from the count of entries and so from the count of associated entries.

Table 2 (cont.)

Entry(E)	Base(B)	Associated Entries(A)	Sum A
chef			
circa			
^{1}clabber	clabber	^{2}B	1
clonal	clone	B	1
collation	collate	B	1
compaction	compact	^{1}B, ^{2}B	2
^{2}composite	compose	B, ^{1}Bite, Bition	3
conclusion	conclu-	Bde, Bsive	2
confide	confide	Bing, Bence, Bential	3
conjuncture	conjunct	^{1}B, ^{2}B, Bion, Bive	4
^{1}consonant	consonant	Bal	1
contentment	content	^{1}B, ^{2}B, ^{3}B, Bed, ^{3}disB, ^{4}disB, disBed	7
counteractive	act	^{1}B, ^{1}Bing, Bion, Bionable, Bivate, Bive, Bivation, Bivity, counterB, inBion, inBive, inBivate, interB, interBant, interBion	15
covenantor	covenant	^{1}B, ^{2}B, Bee	3
^{1}crane			
^{2}creosote	creosote	^{1}B	1
^{1}cropper	crop	^{2}B	1
cruller			
^{1}curry			
dagger			
Darwinism	Darwin	Bian	1
^{1}debauch	debauch	^{2}B, Bee, Bery	3
^{1}decline	decline	^{2}B, Bable, Bation, declension	4
deferment	defer	^{1}B, ^{2}B, Bed, Bable	4
delict			
^{1}den			
depot			
designation	design-	^{1}Bate, ^{2}Bate, Bee	3
detest	detest	Bable, Bation	2
diatropic	trop-	^{3}Bic, Bism, diaBism	3
difficulty	difficult	B	1
disability	ab(i)le	B, Bity, disB, inBity, unB	5
^{1}discredit	credit	B, Bable, ^{2}disB, disBable	4
dismissal	dismiss	B, Bion	2
dissolution	sol-	disBve, disBute	2
divers			
doff			
donna			

Table 2 (cont.)

Entry(E)	Base(B)	Associated Entries(A)	Sum A
^{1}drape	drape	^{2}B, Ber, Bery, unB	4
^{1}earn	earn	Bings	1
^{1}effuse	effuse	Bion, Bive	2
eleven	eleven	Bth	1
ensilage	sil(e)	Bage, Bo, enB	3
epos			
erose			
excursionist	excurs-	Bion, Bive, Bus	3
expand	expan-	Bsion, Bsible, Bsivity, Bsionary	4
exposure	expose	B, Bed	2
extraordinariness	ordinary	^{2}B, Bly, Bness, extraB, extraBly	5
fallacy	fal(l)-	Bacious, ^{1}Bse, Bsehood, Bsification, Bsify, Bsifier Bsity	7
^{2}feeling	feel	^{1}B, unBing	2
ferrous	ferr-	Bic, nonBous	2
filmdom	film	^{2}B	1
firing	fire	^{2}B	1
flagellum	flagell-	Bar, ^{2}Bate	2
flaxy	flax	B, Ben	2
^{3}following	follow	^{2}B	1
forceful	force	^{1}B, ^{2}B, Bed, Bible, enB	5
forlorn			
^{2}frappe			
freshen	fresh	^{1}B, reB, reBen, reBment	4
functionary	function	^{1}B, Bal	2
^{2}gauntlet			
genetics	gen(e)	B, Betic, Beticist	3
giddap			
glade			
glossa	gloss	Bal	1
^{1}hammer	hammer	^{2}B, Bed	2
^{3}harrow	harrow	^{2}B	1
^{2}heather	heath	B	1
helix			
hesitate	hesit-	Bance, Bancy, Bant, Bation, Batingly	5
hominid	homin-	Boid	1
hooklet	hook	^{1}B, ^{2}B, Bed, unB	4
^{1}host			

Table 2 (cont.)

Entry(E)	Base(B)	Associated Entries(A)	Sum A
^{2}hull	hull	^{1}B	1
^{1}icing	ice	^{1}B, ^{2}B	2
impeller	impel(l)	^{1}Bant, ^{2}Bant, B	3
incalculable	calcul-	Bate, Bable, Bability, Bated, Bating, Bation, Bator, inBability	8
inconformity	conform	^{1}B, Bance, Bation, Bity, nonBist, nonBity	6
indemnifier	indemn-	Bify, Bification, Bity	3
individuality	individu-	^{1}Bal, ^{2}Bal, Balism, Balist, Balization, Balize, Bate, Bation	8
^{1}industrial	industry	^{2}Bal, Balism, Balist, B, Balization, Balize	6
^{2}infinitesimal	finit(e)	B, Bude, ^{1}inBsimal, inB, inBude, inBy	6
ingression	ingress	B, Bive	2
^{1}instantly	instant	^{1}B, ^{2}B, Baneous, Ber	4
^{2}intensive			
^{4}intern	intern	^{5}B	1
intramolecular	molecule	B, Bar	2
inviolable	viol-	^{1}Bate, Bation, Bence, Bent inBability, inBate	6
irradiancy	radi-	^{1}Bant, ^{2}Bant, Bance, ^{1}Bate, ^{2}Bate, Bation, irBance, irBant, irBate, irBation	10
ivy			
japanize	Japan	Bization, Bese	2
^{2}joke	joke	^{1}B	1
kersey	kersey	Bmere	1
laic	lay	^{5}B, Bity, Bicism, Bicization, Bicize	5
languid	langu-	Bish, Bor, Borous	3
^{3}latch	latch	^{1}B, ^{2}B	2
^{2}leak	leak	^{1}B, Bage, By	3
^{2}legislative	legislate	B, ^{1}Bive, Bion, Bor, Bure	5
liberality	liberal	^{1}B, ^{2}B, Bism, Bization	4
ligature	ligat-	Bion	1
liminal	limen	B, subliminal	2
^{1}link	link	^{2}B, Bage	2
litigable	litig-	Bant, Bate, Bous	3
localize	local	^{1}B, Bization, Bity	3

Table 2 (cont.)

Entry(E)	Base(B)	Associated Entries(A)	Sum A
^{1}lot	lot(t)	^{2}B, Bery	2
lucky	luck	B, Bily, Biness, unBy, unBily, unBiness	6
lustrate	lust(e)r-	^{2}B, ^{3}B, Bous	3
machinable	machine	^{2}B, Bist, Babilit[illegible]	3
^{1}make	make	^{2}B, Ber, Bing, reB	4
massy	mass	^{2}B, ^{3}B, Bive	3
meanness	mean	^{1}B, ^{2}Bly	2
medico	medic	^{2}B, Bal, Bate, Binal, Bine	5
^{3}mere			
^{2}mistake	mistake	^{1}B, Bable	2
modesty	modest	B, imB	2
mop			
mortify	mort-	Bification	1
mouser	mouse	^{1}B	1
^{2}narrow	narrow	^{1}B, ^{3}B	2
^{4}nigh	nigh	^{2}B, ^{3}B	2
^{1}nonplus	nonplus	^{2}B	1
nosing			
^{1}null	null	^{2}B, Bify, Bification, Bifier, Bity	5
obnoxious	nox-	Bious	1
^{2}officer	office	^{1}Ber	1
opposition	oppose	B, Bable, Bability, Bless, ^{1}Bite, ^{2}Bite, unBed	7
ova	ov-	Bum, Bal, Bate	3
parenteral	enter-	Bal, Bic, Bon	3
partible	part	^{2}B, Bed	2
patriot	patriot	Bic, Bism, unBic	3
pentad			
peremptoriness	peremptory	B, Bly	2
periscope	scope	^{2}B, periBic	2
^{1}personal	person	B, ^{2}Bal, Bality, Balize, Bally, imBal	6
placid			
pleasureless	pleasure	^{1}B, ^{2}B, Bable, Bability, disB	5
poeticize	poet	poem, B, Bry, Bic, Bical, Bicism, Bicalness	7
^{2}police	police	^{1}B, Bman	2
postural	posture	^{1}B, ^{2}B	2
powerful	power	^{1}B, ^{2}B, Bless	3

Table 2 (cont.)

Entry(E)	Base(B)	Associated Entries(A)	Sum A
pregnancy	pregn-	^{2}Bant, ^{1}imBate, ^{2}imBate, ^{1}imBable, ^{2}imBable, imBant imBability	7
prescind			
preterminal	termin-	^{1}Bal, ^{1}Bate, Bus, Bation, Bative, Bator	6
primp			
^{1}problem	problem	^{2}B, Batic	2
promptitude	prompt	^{2}B	1
proration	rate	proB, ^{2}B, Bion	3
^{2}punch	punch	^{3}B	1
purposely	purpose	^{1}B, ^{2}B, Bive	3
pyknic			
readership	read	^{1}B, Ber, Bing	3
recency	recent	B	1
recrystalize	crystal	B, Bize	2
reduction	reduce	B	1
^{1}relish			
^{2}renown	renown	^{1}B, Bed	2
repress	repress	Bed, Bion	2
responsive	respon-	^{1}Bd, ^{1}Bdent, Bder, Bse	4
retral			
revetment	revet	B	1
ridgy	ridge	^{1}B, ^{2}B	2
^{1}rinse	rinse	^{2}B	1
sabra			
sailer	sail	^{2}B	1
saltiness	salt	^{1}B, ^{2}B, ^{3}B, Bed, Bish, By unBed	7
^{1}scoop	scoop	^{2}B	1
^{1}screen	screen	^{2}B, Bing	2
^{1}seat	seat	^{2}B, Bing, unB	3
secularization	secular	^{1}B, Bism, Bity, Bize	4
seller	sell	^{1}B, ^{2}B, sale, sales	4
semitropics	tropic	^{2}B, Bal	2
separation	separ-	^{1}Bate, ^{2}Bate, Bable, Bability inBable, inBability	6
^{3}serpentine			
settler	settle	^{2}B, Bment, unB, unBed	4
^{2}shag	shag	^{1}B, ^{3}B, Bily, Biness, By	5
sharpie	sharp-	Ber	1
shocker	shock	^{4}B, Bing	2
^{5}shy			

Table 2 (cont.)

Entry(E)	Base(B)	Associated Entries(A)	Sum
sigma	sigm-	Boid, Bate	2
silveriness	silver	^{1}B, ^{2}B, By, Bly	4
Slavonian	Slav	B, ^{1}Bic, ^{2}Bonian, Bonic	4
^{1}sole	sole	^{2}sole	1
sou			
spavin			
stet			
stipulation	stipulate	^{1}B	1
stoolie			
^{1}strait	strait	^{2}B, Ben	2
stroma			
^{4}stunt	stunt	^{3}B	1
^{2}subsoil	soil	^{3}B, ^{1}subB	2
supererogation	erogate	superBory	1
supporter	support	^{1}B, ^{2}B, Bable, Bive	4
surveillant	surveil-	Bance	1
swarthiness	swarth	^{3}B, By	2
synonymist	synonym	B, Bize, Bous, By	4
^{1}target	target	^{2}B	1
^{1}temporal	tempor-	Bality, Balize, Bize, Bization, Barily, Bariness, Bary	7
tertoma			
^{2}testimonial	test-	Bimony, Bament	2
^{4}that			
^{1}thinking	think	^{1}B, ^{2}B, ^{3}B, ^{2}Bing, Bable, ^{1}thought, ^{2}thought, thoughtful, thoughtless	9
threadiness	thread	^{1}B, ^{2}B, By, Bless, unB	5
thug	thug(g)	Bee	1
townee	town	B	1
transgress	transgress	Bion	1
trash	trash	By, Biness	2
^{1}trend	trend	^{2}B	1
tripe			
turgescence	turg-	Bid, Bescent	2
^{3}turtle	turtle	^{2}B	1
ultima			
unbind	bind	B, unbound, bound	3
^{2}unequal	equal	^{1}B, ^{2}B, ^{3}B, ^{1}unB, Bity, Bize, Bization, Bizer, Bly, inBity	10
unitarian			

Table 2 (cont.)

Entry(E)	Base(B)	Associated Entries(A)	Sum A
unprofessed	profess	B, Bed	2
untrue	true	^{1}B, ^{2}B, Bly, Bism, unBth, unBthful, Bth, Bthful	8
utilitarian	utilit-	By, Barianism	2
validity	valid	B, Bate, Bation, inB, inBate	5
variety	vary	B, Bous, Betal, ^{1}Bable, ^{1}Bant, Bance, Bative, Bational, inBant, inBance, inBable, inBability	12
^{1}veil	veil	Bed	1
verticality	vertic-	Bal	1
vicarate	vicar-	B, Bage, Bial, Biate	4
vindication	vindicate	B, Bion, Bory, vindicable	4
viscous	visco(u)s-	Bity	1
vocalist	vocal	B, Bize, Bization	3
wageless	wage	^{2}B	1
^{1}wallow	wallow	^{2}B	1
^{1}warning	warn	B, ^{2}Bing	2
^{1}whin			
^{1}widow	widow	^{2}B, Ber, Bhood	3
worldly	world	B, Bling, Bliness, unBly, unBliness	5
wright			
zonation	zone	^{1}B, ^{2}B, Bal, Bate, Bary, reB	6

Sum E (over pages) = 266

Sum A (over pages) = 617

The associated entries (A) of Table 2 probably reflect some entries that do not qualify as purely-linguistic extensions of the base form in that they introduce a nuance of meaning not detectable from the base form and its extension consonant with provisions of an appropriate linguistic rule. Probably offsetting these entries are others not detected in Webster's due to a changed prefix that did not come to mind during the search for associated entries. These offsetting possibilities considered, we will consider Sum A an unbiased value.[1]

The use of Category 5 (for entries that repeat a Category 4 base form) insures that the sum of Category 4 entries must equal the sum of the Category 4 base forms. Sum E = Sum B = 266. Since base forms do not recur in the Category 4 sample, neither should any entry of (Sum A + Sum E) occur more than once.

The domain of a base form B_i is $A_i + E_i$. The sum of these domains is (Sum A + Sum E) = 619 + 266 = 885. The proportion Sum B/(Sum A + Sum E)-- 266/885 = .30--can be used to estimate how many base forms are reflected in estimated Webster's Category 4 entries (36K, see Table 1). This estimate is .3(36K) = approximately 11K. We might interpret this value as asserting that the estimated 36K Category 4 entries of Webster's reflect approximately 11K base concepts, exclusive of linguistic concepts. Perhaps these concepts give way to 36K concepts in consequence of acquisition of a non-inordinate number of linguistic concepts--e.g., 500--or rules. One approach to specifying these rules (and ordering them for usefulness) would be to elaborate on the Table 2 sample and to analyze the sample for productive rules of the base extension type.

[1]Extensions of base forms typically are orthographically regular in Table 2. A few irregular extensions were allowed. For example, the domain of the base form think includes the entries THINK and THOUGHT and their extensions. This practice was followed only in the case of a few common irregularities.

Table 3.

Non-Category 4 Entries in the Webster's Sample of 518 Entries

Category	Entries			
1	aecium	[2]drivel	Marathi	Sarum
	[1]agnate	duad	marquisette	[1]scaphoid
	ambry	Duralumin	mattery	schlieren
	appanage	eclogue	Memnon	[2]scutch
	archicarp	electroanalysis	metalographic	sejant
	Armistice Day	enantiomorph	Middle Greek	shoulder girdle
	autotomize	epexegesis	nematocidal	six-o-six
	benzyl	essoin	nociceptive	small ale
	bibliofilm	euglena	oligosaccharide	snowblink
	billbug	evonymus	oogonium	[2]spencer
	[1]birds-eye	facia	Ordovician	sporozoite
	blackleg	[4]fly	overtrade	[2]spunk
	blesbok	frowsty	[5]pack	suberization
	bonnyclabber	futtock	palet	suderific
	[2]Brythonic	galliverous	panetela	summa cum laude
	[4]bull	gerardia	paradisiacal	symmetallism
	burning bush	Graeae	pearl danio	Tammuz
	butylate	greenbrier	Pelagian	tin can
	canoness	grosz	pirarucu	tmesis
	cellarette	guide word	Polyhymnia	tory-rory
	[1]chamfer	hansel	posada	tropaeolum
	[3]choir	hirple	procession of the equinoxes	trumpet flower
	coalfish	hylozoism		tuckerbag
	coelenteron	Igraine	pro forma	twopence
	colza	imide	pseudepigraphy	Venusberg
	[2]commerce	impresa	quadrivium	vorlage
	conurbation	inquiline	quinoidine	wave front
	copeped	Jumada	raceme	weaverbird
	corvine	juvenilia	radula	wheatear
	[1]culm	knobkerrie	reichsmark	[3]wind
	Cybele	kurus	rochet	withe rod
	dextrogyrate	Leto	Roman collar	wood spirit
	[2]douse	[5]long	roset	yellowhammer
2	airworthiness	bank paper	broad spectrum	chili sauce
	allspice	beachcomber	cardiorespiratory	clear-eyed
	analogue	bedbug	carrying charge	Corn Law
	apropos of	bobwhite	caster oil	[2]dingdong
	au jus	brassbound	catnip	endothermic
	backlash	brewer's yeast	cerebral hemisphere	[1]farewell

Table 3 (cont.)

Category	Entries			
2	[1]faultfinding	millstone	poor box	stalemate
	fiddler crab	minuteman	protohistoric	staphylococcal
	floor furnace	monaxial	public speaking	stationmaster
	foulmouthed	monograph	quasi-legislative	steering wheel
	garden cress	multicellular	ramshorn	streptolycin
	go-between	musical saw	rate payer	swineherd
	goose egg	myriameter	reserve bank	table wine
	gynecocracy	Neanderthal man	round worm	taxiway
	haircloth	neuromuscular	ruffed grouse	tie-in
	hazelnut	nut-brown	run-of-the-mill	tollman
	hen party	octagon	sand myrtle	traffic engineer
	hieroglyphic	orthograde	sawlog	tricostate
	holdback	Our Father	self-explaining	underdeveloped
	hypomania	pass out	shepherd dog	upwind
	isoline	petticoat	singsong	water ballet
	jet stream	philanthropic	ski pole	well-grounded
	Law of Moses	photocopy	slipslop	white pine
	magnetic flux	physiological	snail-paced	winterberry
	malnutrition	pincushion	sowbelly	xerothermic
	manhunt	plaster of paris	spinal column	you've
	metric system	plumb rule	squinting construction	
3	-chrome	-kinesis	socio-	tel(e)-
	embry(o)-	piezo-	-somic	therm(o)-
	grapho-	rhod(o)-	splen(o)-	top(o)-
	hyal(o)-			
5	refresh			

The word formation rules to be learned in primary education select and order morphemes--base morphemes and bound morphemes such as prefixes, suffixes, combining forms, and, occasionally, infixes. These bound morphemes are estimated in Table 1 to be approximately 2K in number. A supplemental inspection of Webster's shows some Category 3 entries to be repetitive. Let us imagine that the unique Category 3 entries contained in Webster's are on the order of 1K. A casual review of Webster's suggests that a preponderence of its affixes will be addressed by word formation instruction during primary education, but that less than a preponderence of the combining forms will be. (The latter probably name concepts that will be treated in advanced science and technology courses given during postprimary education.) Perhaps no more than 500 of the unique Category 3 entries of Webster's will be addressed during primary education. Let us call these base concept names of the Category 3 type.

The picture that emerges in consequence of sampling and speculation is that the child will leave primary education understanding some 11K base concepts of the Category 4 type and some 500 base concepts of the Category 3 type. (The former will include linguistic concepts for which there are entries in Webster's; some representative base forms for such entries are noun, apostrophe, sentence, and tense.) If the child also is able to apply some 500 word formation rules to these base concepts, then his understanding will be extended to some 25K related "higher level" concepts, hypothetically on a purely linguistic basis. (This view assumes that the inflationary effects of synonymity upon estimates of extended concepts of Webster's will be offset by counteractive multiple-meaning effects not taken into consideration above.)

Later efforts referencing to specification of a lexicon for primary education would have to evaluate whether it is warranted to use Webster's as a source reference for clues to the base concepts of interest to primary education. Such efforts also would need to enunciate a more defensible basis (than intuition) for distinguishing between primary and postprimary base concepts and attendant concept names. Finally, these efforts would need to determine a definitive basis for dealing with synonymity and multiple meanings. Assuming that these matters can be dealt with in a satisfactory way, it might make good sense to wed efforts to specify lexicon to the base concept notion--in the sense that the concept is used above, rather than in the sense of the lowest level of a grand conceptual hierarchy encompassing primary education.

The child enters primary education understanding an appreciable number--although probably much less than a preponderence--of the estimated 11K Category 4 and .5K Category 3 base concepts that we assert he will take from primary education. Perhaps the search for a lexicon for primary education is grounded on a search for something on the order of 9K base concepts.

The foregoing remarks are intended to advance an effort to formulate a comprehensive Communication Skills program for primary education. Earliest views on such an effort have given specification of lexicon a central--almost propaedeutic--role in the effort. While lexical specification continues to appear central and to require early treatment, it now appears also tenable that such work should give prominence to the referents--and particularly the base concept referents--that entries of the lexicon name. The base concept has been defined above in such a way that, if employed in lexical specification, it need presume little if anything concerning how concepts of the referential domains are organized. Such matters belong to the other facets of the overall effort (which might be able to use the base concepts of lexical specification as a point of departure to questions of organization).